

Objekt, Prozess und Kontext: Gedächtnisorganisationen und Semantische Technologien

Prof. Dr. Stefan Gradmann

Präsident der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis (DGI), Humboldt-Universität zu Berlin

Transkript des Vortrags vom 17.11.2011

Vorhin wurde behauptet, dass Europeana auf dieser Veranstaltung bislang eher als Phantom zu gegen gewesen sei. Ich bin damit in gewisser Hinsicht die Inkarnation dieses Phantoms und möchte zu Ihnen über einen relativ abstrakten Gegenstand sprechen. Der Titel meines Vortrags lautet "Objekt, Prozess und Kontext". Es geht mir darum, welche Potenziale semantische Technologien für Gedächtnisorganisation eröffnen - und natürlich geschieht das am Beispiel von Europeana.

Hier meine Gliederung: Ich möchte beginnen mit meiner Sicht dessen, was traditionellerweise die Objektperspektive in Museen, Archiven und Bibliotheken gewesen ist, in Absetzung zur Prozessperspektive, die im zweiten Teil meines Vortrags eine größere Rolle spielt. Ich komme dann zu technischen Aspekten, denn wenn wir über semantische Technologien sprechen, ist eine Verständigung darüber erforderlich, was das eigentlich ist. Ich gehe also auf RDF ein, das Resource Description Framework, ich schlage von dort den Bogen zum Europeana Data Model und spreche dann darüber, wozu das Ganze eigentlich gut ist. Es geht letztlich um einen Ansatz für die Kontextualisierung und die hier aufgeführte Gleichung „Wissen = Information + Kontext“ ist konstitutiv für den dritten Teil des Vortrags. Schließlich komme ich darauf, wie wir in Zukunft auf dieser neuen technischen Basis Objekt- und Prozesssicht in Gedächtniseinrichtungen anders aufeinander beziehen können. Dies bringt mich abschließend zu neuen Fragen an und neuen Chancen für Gedächtnisorganisationen.

Auf meiner ersten Folie sehen Sie die gerade untergegangene Welt der Bibliotheken. Es ist noch gar nicht lange her, im Jahr 2005, als dies Bild des große Zettelkatalogs in der Sterling Memorial Library aufgenommen wurde. Kataloge dieser Art waren sozusagen die Herzen von Bibliotheken. Sie verzeichneten Objekte. In dem Wort Katalog steckt das Verzeichnis: *καταλογος* ist das "Verzeichnis", die "Liste", ein Warenangebot, das Verb dazu, *καταλεγειν*, kann man übersetzen mit "aufzählen" oder "in Listen eintragen". Es geht dabei immer um etwas Abgeschlossenes, einen Bestand, etwas, was prinzipiell lokalisierbar ist, in einer Bibliothek, abgeschlossen und aufzählbar und das bedeutet auch die Entitäten in diesem Bestand müssen gleichermaßen ‚listenfähig‘ sein, sie müssen in sich abgeschlossene Entitäten sein, diskrete Objekte. Diese Objekte und ihre Attribute haben Bibliothekare jahrzehntelang auf Zetteln in ihren Katalogen beschrieben. Sie tun es heute mit nicht wirklich ansprechenden MARC-Datensätzen, wie auf der folgenden Folie zu sehen. In der Sache tun sie damit aber nichts Neues, sie referenzieren unverändert etwas, was sie als ‚Bestand‘ begreifen. Dieser Bestand ist konstituiert durch Objekte, die in sich geschlossen monolithisch gedacht sind. Indem sie die Bestandsobjekte in Katalogen beschreiben tun sie letztlich gleich, ob sie nun Bücher referenzieren oder PDFs oder genuin digitale Objekte, immer das Gleiche, wie in der nachstehenden Abbildung dargestellt:

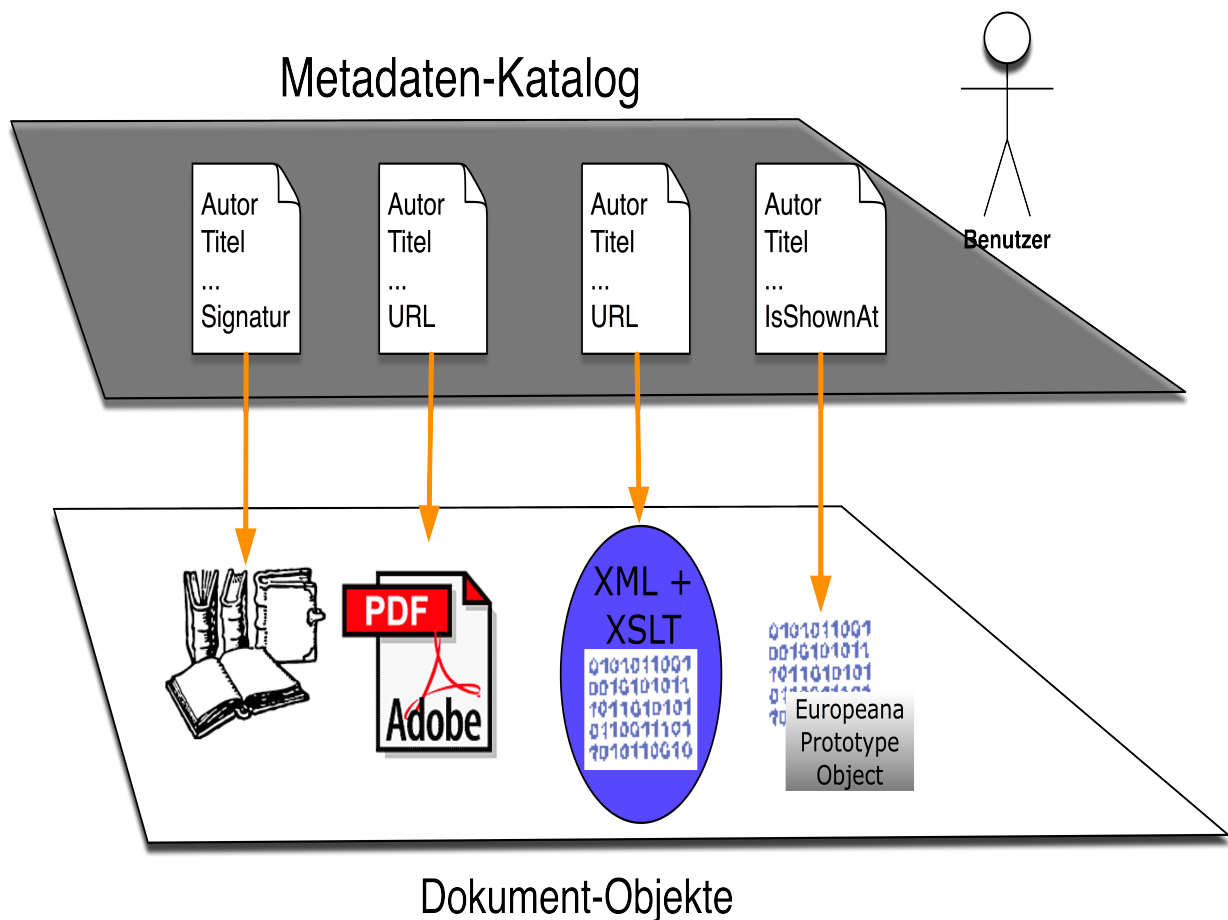


Abbildung 1: Metadaten und Objekte in Bibliothekskatalogen

Sie erstellen einen Metadatensatz, dieser enthält einen Zeiger auf das Objekt und dieser vermittelt den Zugriff auf das Objekt selbst, der eben über diese Katalog gewährleistet wird. Der Benutzer bewegt sich immer durch die vermittelnde Schicht der Katalogdaten und erhält aus schließlich darüber den Zugriff auf die Objekte selbst. Es handelt sich um ein völlig statisches System für die Verwaltung von Informationscontainern. Diese objektgeleitete, statische Sicht ist in Bibliotheken immer prägend gewesen. In Museumsordnungen ist die Sicht – wenngleich immer noch objektkonstituiert – ein wenig anders: auch die Sammlungen von Museen und die Art und Weise, wie sie präsentiert werden, entwickeln sich meist sehr stark vom Einzelobjekt her, aber das Objekt ist sehr viel stärker kontextualisiert. Museen verwenden sehr viel mehr Zeit darauf, Kontext für Objekte herzustellen. Bibliothekare tun dies in der Regel nicht. Und schließlich ganz anders ist die Perspektive in den Archiven. In den Archiven gibt es eigentlich, sehr vereinfacht gesprochen, gar keine Objekte ohne Prozesse und ohne Kontext, d.h. der Provenienzkontext, der in das Objekt eingeschrieben ist, konstituiert ihn eigentlich erst: die Perspektive ist im Grunde gegenüber den objektgetriebenen Bibliotheksperspektive genau verkehrt.

Hier nun meine erste These: Die objektgetriebene bibliothekarische Logik hat bislang eigentlich alle Kulturportale dominiert, bis hin zu dem, was sie heute von Europeana im Netz sehen können.

Und nun, zum nächsten Teil überleitend, die zweite These gleich dazu: Ein RDF-graphenbasiertes Modell, wie das Europeana Data Model (EDM), kann dies grundlegend ändern. Für das Verständnis dieser These allerdings ist es nunmehr erforderlich, den Hintergrund des Attributs

„RDF-Graphenbasiert“ darzulegen – und damit wird die Darstellung notgedrungen technisch. Es folgt der knappe Versuch, darzulegen, was das World Wide Web technisch eigentlich ist, wie es sich entwickelt hat und warum wir derzeit von einem Semantic oder Linked Data Web sprechen.

Das in der folgenden Abbildung dargestellte Dokument ist sozusagen die Geburtsurkunde des WWW.

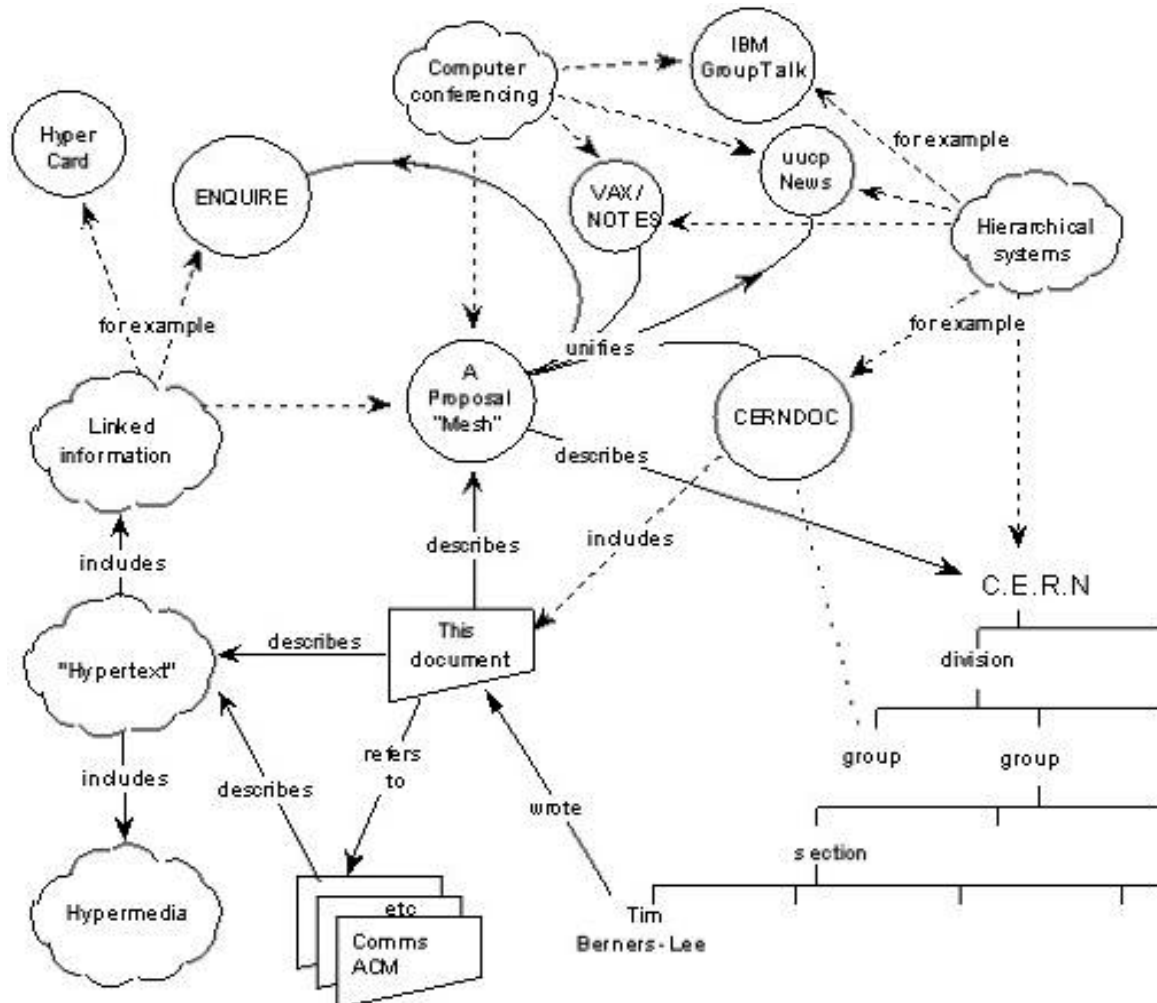
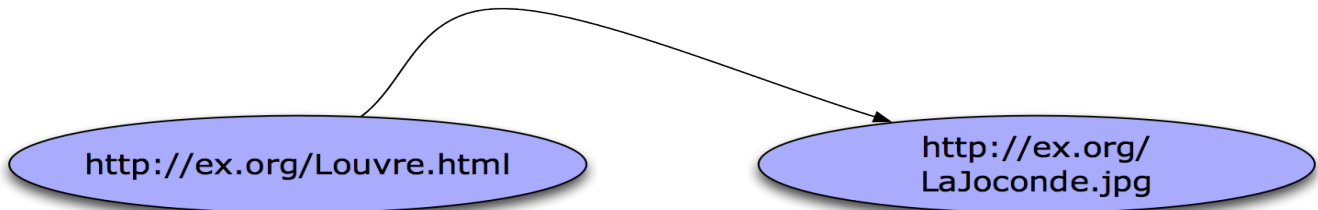


Abbildung 2: Information Management - a Proposal

Es handelt sich um den legendären Vorschlag für ein Informations-Management-System, den Tim Berners-Lee 1989 seinem Arbeitgeber am CERN vorgelegt hat. Diese erste Version des WWW war ein System von Dokumenten, die untereinander verbunden waren, eine gigantische Hypertext-Anwendung. In diesem Bild gibt es, Dokumentressourcen und Verbindungen zwischen ihnen. Diese Verbindungen sind übrigens in dem ersten Vorschlag von Berners-Lee interessanterweise sogar schon typisiert. Gegenüber diesem Bild ist das Linked Data Web eine Erweiterung im doppelten Sinne, nämlich einerseits im Sinne der Syntax, der Art und Weise Aussagen über Objekte im Web zu ermöglichen, und im Sinne des Umfangs, denn das ursprüngliche Web in der Vision von Berners-Lee war ein Web von Dokumenten und das Linked Data Web erweitert dieses Paradigma indem es so gut wie alles in der Welt im Web repräsentierbar macht.

Die folgende Abbildung illustriert die einfache Syntax der einfachen HTTP-Graphen im WWW der ersten Generation:



```
<a href="http://ex.org/LaJoconde.jpg">b</a>
```

Abbildung 3: Ein HTTP-Graph

Der Graph setzt zwei Web-Ressourcen miteinander in Verbindung, nämlich auf der linken Seite Louvre.html und auf der rechten LaJoconde.jpg. Für uns Menschen ist damit schon eine Menge Information gegeben, denn wir können Vermutungen darüber anstellen, was für Arten von Entitäten Louvre.html und LaJoconde.jpg sind. Wir wissen vielleicht, dass La Joconde der französische Name für das Bild ist, das wir „Mona Lisa“ nennen, wir wissen, dass „Louvre“ ein Museum ist. Wir wissen, dass beides französische Wörter sind. Wir können also einen kulturellen Kontext inferieren. Eine Maschine kann das alles nicht. Sie kann keine Vermutung darüber anstellen, was für einer Klasse von Objekten diese beiden Ressourcen eigentlich angehören.

Gleiches trifft auf die Beziehung zwischen beiden zu: wir können diese Beziehung inferieren: Wir wissen, dass Bilder in Museen hängen, also nehmen wir an, dass dies letztlich die Aussage dieses Graphen ist: La Joconde hängt im Louvre. Auch hier wieder hat die Maschine keine Chance, zu einem solchen Schluss zu kommen, also müssen wir es ihr den Zusammenhang zwischen den beiden Ressourcen explizit machen - und genau das leistet RDF. RDF, das Resource Description Framework, typisiert Links und ermöglicht die im folgenden Beispiel illustrierte Aus

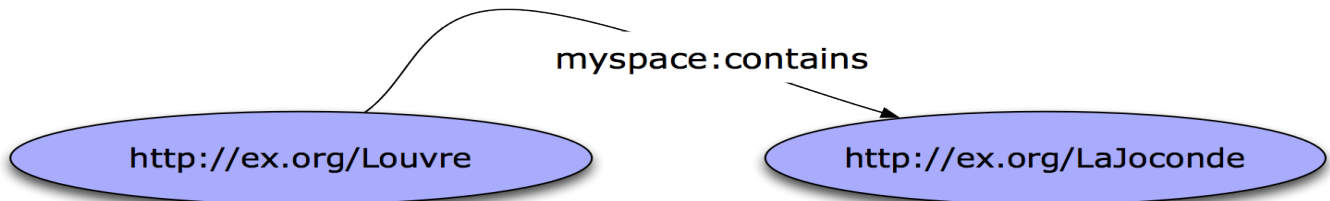


Abbildung 4: Ein RDF-Tripel

- die eine Ressource enthält die andere, oder, allgemeiner gesprochen, wir erweitern die Syntax dessen, was man im Web aussagen kann dergestalt, dass wir mit dreiwertigen Sätzen, sogenannten Tripeln, die immer ein Subjekt, ein Prädikat und ein Objekt enthalten (oder eine Ressource, eine Property, einen Value), damit Aussagen machen können über Ressourcen im Web. Bei diesen Tripeln handelt es sich in gewisser Hinsicht um die Atome dessen, was seit ein gen Jahren 'Semantic Web' heisst. Zum Semantic Web oder zum Linked Data Web wird eine Aggregation solcher Tripel dadurch, dass wir diese einfachen Aussagen mit einem Grammatikmodell kombinieren, mit einer Schemasprache (RDFS), welche weitere Konstruktoren enthält sowie Elemente, die es mir erlauben, Sätze in Klassen zu organisieren, Hierarchien von Klassen und Eigenschaften zu modellieren, und innerhalb dieser Hierarchien Eigenschaften zu vererben. Damit erhalten wir Unterstützung für einfache Schlussysteme, auf die ich am Schluss dieses Vortrags dann auch noch einmal praktisch kommen werde. Damit werden Aggregationen von Tripeln organisierbar, denn von diesen kleinen Sätzen über Dinge im Web gibt es eben viele, viele Milliarden, und mit RDFS kann ich diesen Strukturen einziehen, kann sie modellieren und kann daraus leichtgewichtige, sogenannte Ontologien machen.

'Ontologien' ist ein weiteres Wort, das die Informatiker von den Philosophen gekapert haben. Bei den Informatikern bedeutet es einfach, dass Wissen über einen Ausschnitt von Welt

mit Hilfe solcher RDF-Statements und der dazu gehörigen RDFS-Grammatik modellierbar ist.

Soweit die syntaktische Erweiterung des Web. Die zweite Erweiterung betrifft den Scope, den Erstreckungsbereich des Web. Das Web der ersten Generation war konzipiert als ein Web von Informationsressourcen, die genuin nur im Web existierten: die Dokumente. Die Erweiterung des Web, die wir momentan erleben, macht es möglich, beliebige Dinge auf der Welt - in der Terminologie des W3C so genannte 'Non-Information-Resources' - im Web zu repräsentieren. Eine Illustration dieser Erweiterung liefert die nachstehende Abbildung:

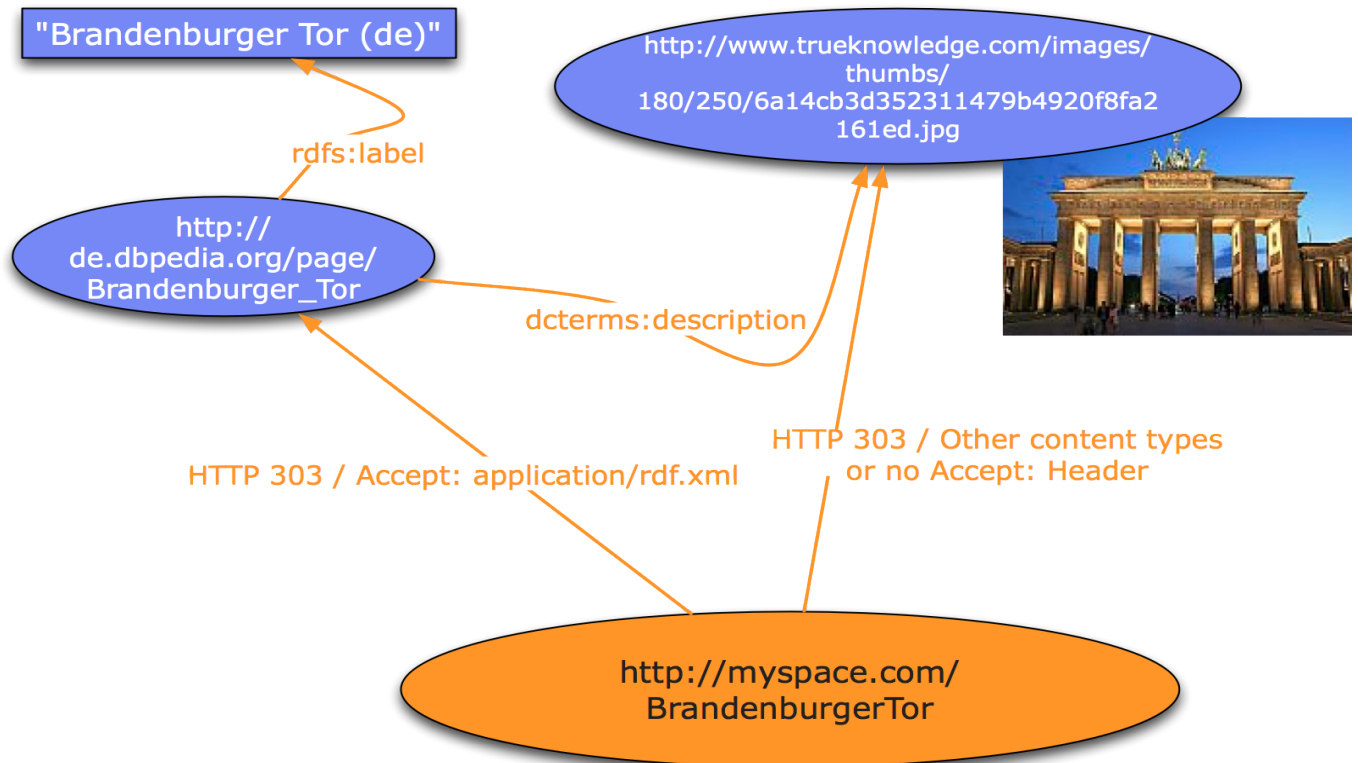


Abbildung 5: Eine Non-Information Resource im WWW

Die orange eingefärbte Ellipse ist hierbei eine Repräsentation eines wirklichen Objektes in der Welt, nicht weit von hier, des Brandenburger Tors. Diese Repräsentation als Webressource ist wieder verbunden mit anderen Ressourcen im Web zwei Triple-Statements, nämlich nach links mit einer Ressource namens dbpedia.org, der RDF-Version der Wikipedia, welche die Konzeptsicht des Objektes enthält auch außerdem noch einen Link zu einer natürlichsprachigen Benennung, und alternativ ist eine Verbindung herstellbar zwischen dieser Repräsentationsressource und einem Bild, rechts oben. Zwischen dem Bild und der Konzeptressource besteht eine Verbindung auch wieder im Sinne von Dcterms:Description: das eine beschreibt das andere. Wieder ein einfaches Set von Tripel-Statements, die in diesem Fall verwendet werden können, um Aussagen über einen Teil der 'wirklichen' Welt außerhalb des Web zu ermöglichen, der hier im Web repräsentierbar wird.

Diese beiden Erweiterungen des Web sind die Basis für in Europeana seit zweieinhalb Jahren erarbeitete Europeana Data Model (EDM). Das Europeana Data Model ist ebenfalls eine Ontologie, die Aussagen möglich machen soll über kulturelle Artefakte aller Art als Teil von Webapplikationen. Angedeutet ist dies im nachstehend abgebildeten Klassenmodell des EDM, in dem die schon erwähnte Unterscheidung zwischen Information Ressource und Non-Information Ressource an der Spitze der Klassenhierarchie konstitutiv ist:

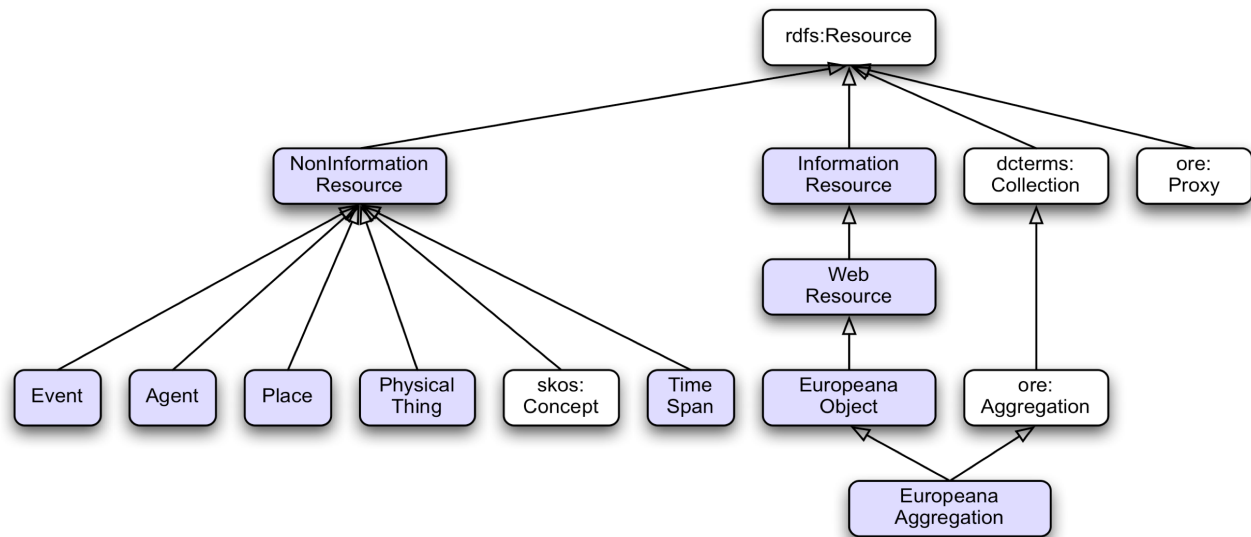


Abbildung 6: Das Klassenmodell des EDM

Unterhalb der 'Wurzelklasse' `rdfs:Resource` befinden sich dabei auf der einen Seite links die Dinge, die in der wirklichen Welt als Kontextinformationen für kulturelle Artefakte existieren und in Europeana modelliert werden müssen wie Ereignisse, Agenten, Orte, physische Dinge oder auch Konzepte oder Zeitspannen. Auf der rechten Seite steht der Begriff Europeana Aggregation. Das ist die Beschreibung der kulturellen Artefakte selbst in einer gleich zu erläuternden Syntax. Entsprechend diesem Klassenmodell gibt es in dieser Ontologie auch ein Eigenschaftensmodell, mit dem die Beziehungen zwischen Instanzen dieser unterschiedlichen Klassen in dem Datenmodell organisieren werden, so dass damit kulturelle Artefakte aller Art repräsentiert werden können.

Das folgende Beispiel mag die Verwendung dieses Modells illustrieren. Nachstehend ist eine traditionelle Repräsentation einer Ausgabe von Flauberts *Fleurs du Mal* im Format UNIMARC zu sehen:

```

000 nam 22 450
001FRBNF373670350000003
009http://catalogue.bnf.fr/ark:/12148/cb37367035f
039 $oGEA$a000288182
100 $a19920409d1857 m y0frey50 ba
1010 $afre
102 $aFR
105 $a||||z 00|||
106 $ar
2001 $a^Les %o%fleurs du mal$bTexte imprimé$fpar Charles Baudelaire
210 $aParis$cPoulet-Malassis et De Broise$d1857
215 $a248 p.$d19 cm
  
```

676 \$a841.8\$v22
 686 \$a840\$2Cadre de classement de la Bibliographie nationale française
 700 |\$311890582\$aBaudelaire\$bCharles\$4070
 801 0\$aFR\$bBNF\$c19920409\$gAFNOR\$2intermrc

Daraus wird in der HTML-Darstellung der Digitalen Bibliothek Gallica das folgende Bild:

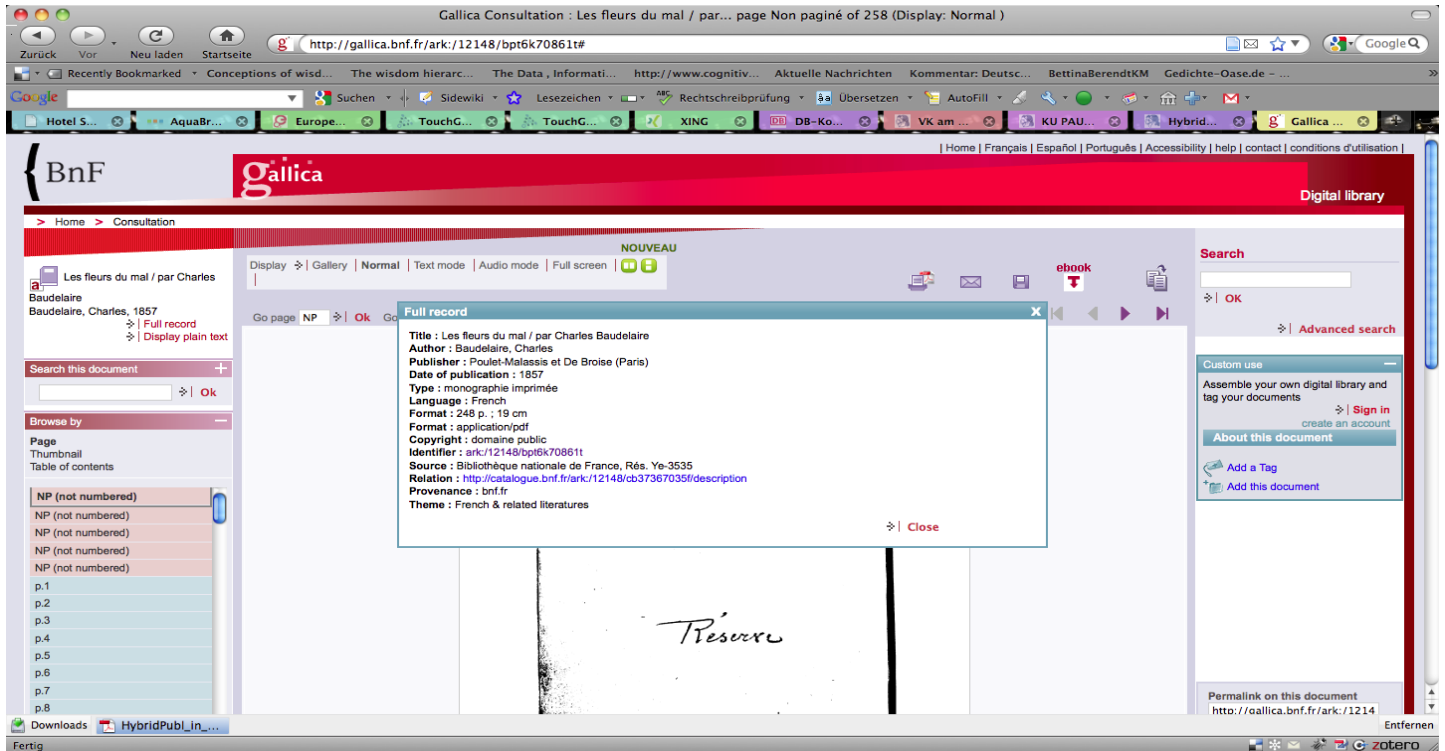


Abbildung 7: Les Fleurs du Mal in Gallica

Und die nachstehende Abbildung illustriert die Modellierung derselben Metadaten-Attribute in EDM:

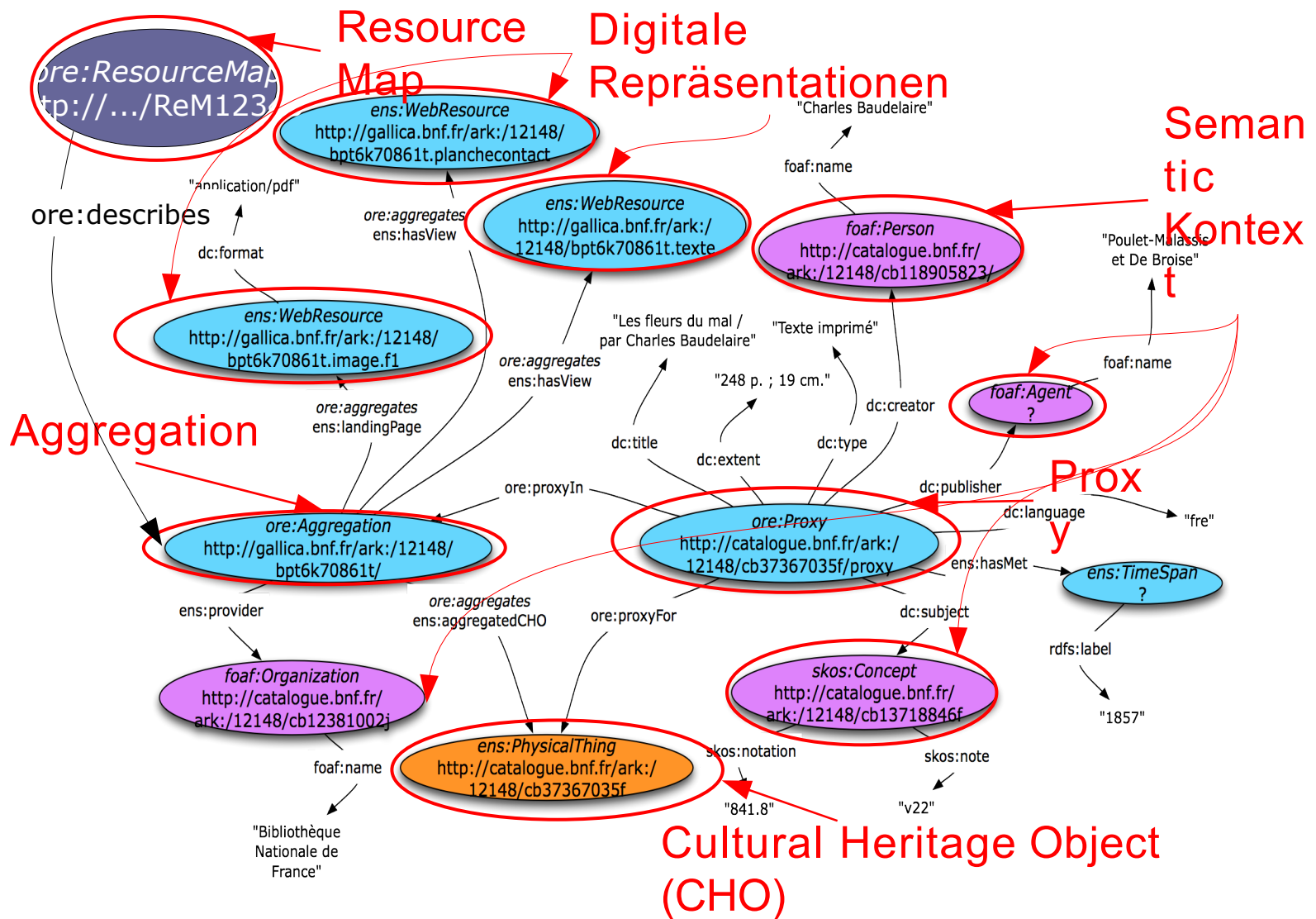


Abbildung 8: Les Fleurs du Mal in EDM

Das orange eingefärbte wirkliche, zu repräsentierende Cultural Heritage Object wäre als Teil der EDM-Aggregation durch ein Proxy-Element repräsentiert, das für eine bestimmte Sicht dieses Objektes steht und dann beschreibende Metadaten und Kontextdaten bindet. Dieses Objekt hat digitale Repräsentationen (Thumbnails, Bilder) und es hat semantischen Kontext, der durch die violett eingefärbten Knoten dargestellt ist: das Objekt ist verbunden mit Personen, mit Konzepten, mit Organisationen. All dies wird zusammengehalten durch einen zentralen Knoten, die sogenannte Aggregation, und diese wiederum wird beschrieben durch eine sogenannte Resource Map. Das entscheidende Charakteristikum dieses Modellierungsverfahrens ist seine Graphenbasiertheit: es handelt sich nicht mehr um einen Datensatz, sondern um eine Aggregation von Tripeln. Dadurch wird es im Prinzip beliebig erweiterbar, dynamisch und verlinkbar.

Hier nun ein zweites Beispiel einer Modellierung in EDM:

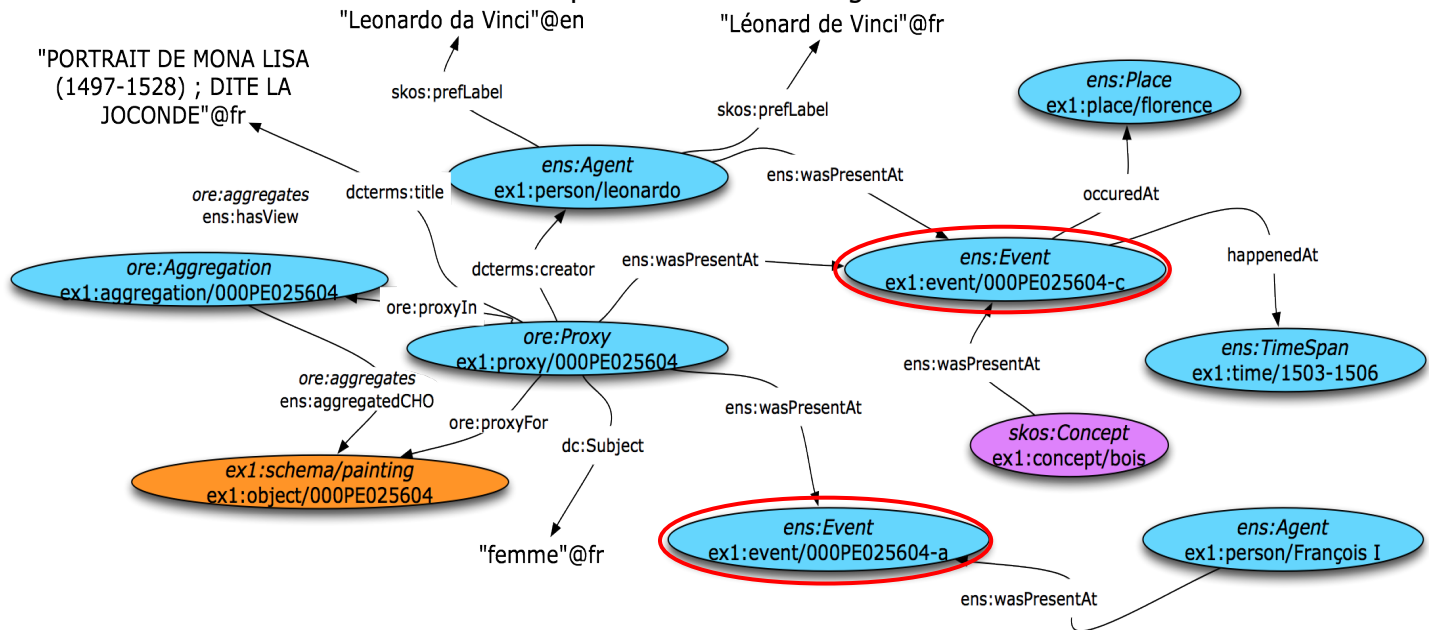


Abbildung 9: Zwei Ereignisbezüge der "Mona Lisa"

HAuch hier handelt es sich um eine Modellierung des Gemäldes der Mona Lisa. Dieses ist hier kontextualisiert, d.h. es wird in Verbindung gebracht mit dem 'Agent' Leonardo als creator, es wird lokalisiert und in der Zeit kontextualisieren, als Teil einer Periode deklariert, und schließlich kann es mit Entitäten aus Wissensorganisationssystemen, hier einem Konzept aus einem SKOS Thesaurus, in Verbindung gebracht werden.

Das Modell muss aber nicht notwendig die Objektsicht einnehmen, sondern eben auch eine prozessbezogene Modellierung vornehmen. In unserem Beispiel war das Gemälde Teil eines Schaffensprozesses, eines `ens:Event` und wurde in einem zweiten Ereignis von König François I. nach Frankreich verbracht. Wir erhalten damit ein neues Potenzial für Gedächtniseinrichtungen: wir können Objekte und ihre Eigenschaften kontextualisieren, wir können das aus einer primären Objektsicht tun oder primär von den Relationen her denken. Dieses relationenorientierte Denken ist etwas, zumindest aus bibliothekarischer Sicht, völlig Neuartiges. Im Prinzip werden in diesem Ansatz die Prozesssicht und die Objektsicht prinzipiell gleichwertig. Das kann die Gewichte in Kultureinrichtungen ganz erheblich verschieben - allerdings nur, wenn tatsächlich alle ihre Daten in einem solchen Modell auch frei verfügbar machen.

Einige von Ihnen werden wissen, dass wir in Europeana eine durchaus noch kontroverse Diskussion haben, in der weniger die Bibliothekare das Problem sind als die Vertreter aus Museen und Archiven, die Schwierigkeiten damit haben, ihre Daten in einem solchen Modell frei verfügbar und frei verlinkbar zu machen. Das ist jedoch unbedingte Voraussetzung dafür, dass das, was ich Ihnen hier zu schildern versuche, sein Potenzial tatsächlich entfaltet, und das muss keinesfalls ausschließen, dass auch dauerhaft erkennbar bleibt, von wem diese Daten kommen, dass eine Provenienzzuschreibung durchaus erfolgen kann.

Was kann man nun mit diesem Modellierungsansatz tun? Wir können – wie schon gesagt – mit diesem Modell Objekte kontextualisieren, und damit schaffen wir letztlich Maschinen für die Wissensgenerierung. Das ist im Kern, was Europeana aus meiner Sicht werden sollte. Das geschieht, indem wir diese einfache Grundfigur des EDM, die Kombination aus Physical Thing, Aggregation und Proxy kontextualisieren, wie in der nachstehenden Abbildung angedeutet:

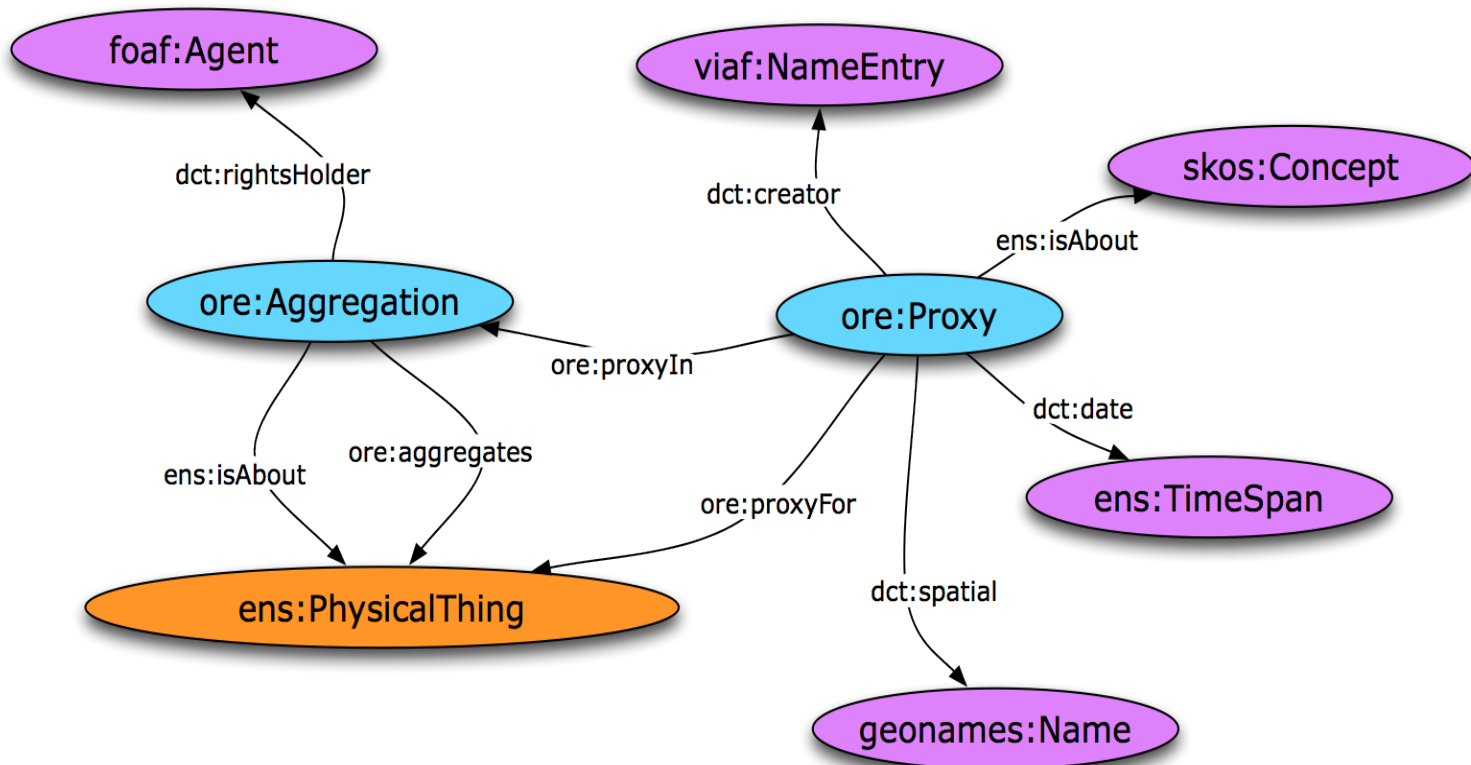
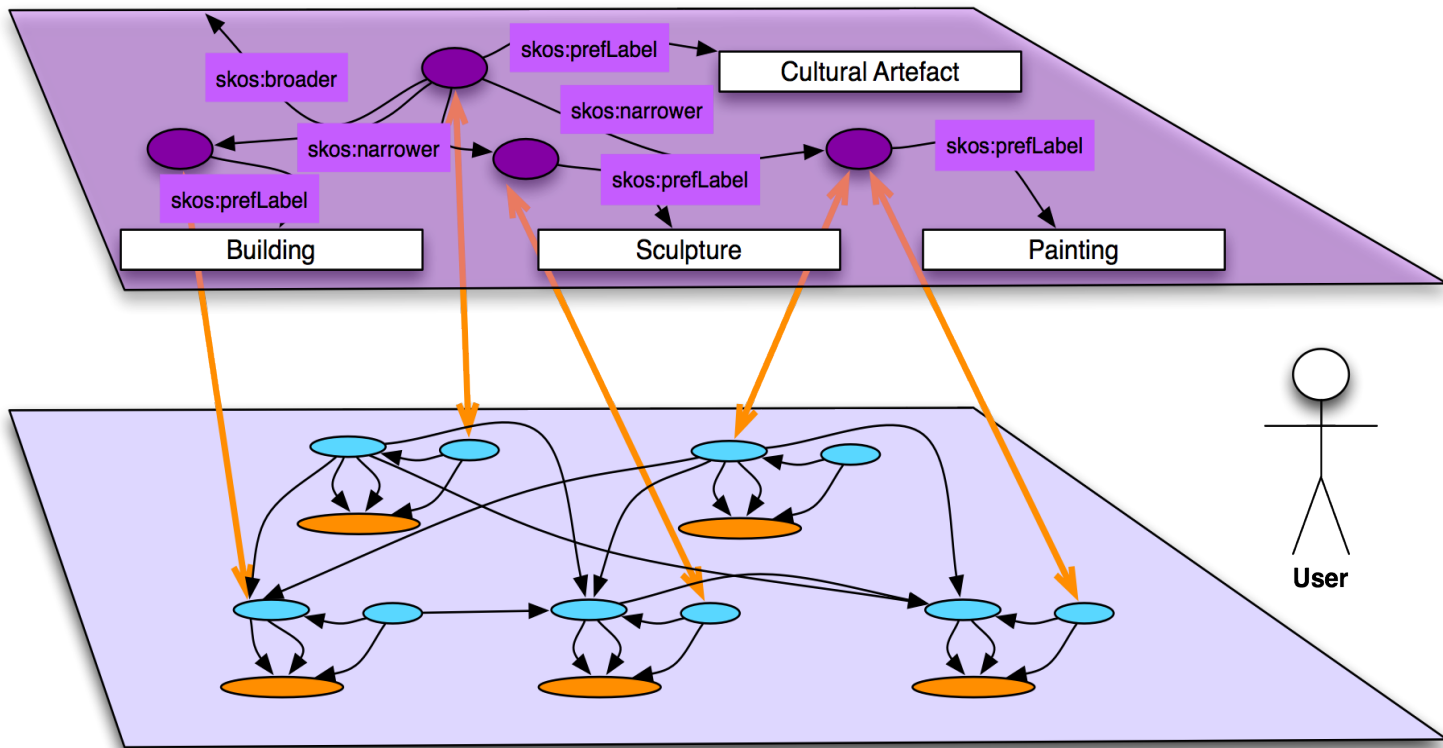


Abbildung 10: Aggregation und Kontext

Solcherart wird die Ebene der vernetzten Objektrepräsentationen Teil eines Bildes, das wir schon sehr früh in Europeana als Vision gezeichnet hatten, und dessen Informationsarchitektur sich grundlegend von dem eingangs gezeigten Bild der bibliothekarischen Metadaten mit samt ihren 'Zeigern' auf die Objekte unterscheidet:

Semantic Network



Networked Object Representations

Abbildung 11: Das Zwei-Ebenen-Modell der Europeana

Es ist dies ein wesentlich interaktiveres Modell, in dem die Objektrepräsentationen, die auf der unteren Schicht zu sehen sind, und die im EDM modelliert sind, untereinander vernetzt und verlinkt sind, aber eben auch verlinkt sind mit der Schicht darüber, der semantischen Schicht, wodurch sie kontextualisierbar werden und damit zu einem Teil von Wissensräumen. Dieses Modell kann man nun seinerseits benutzen, um die Verständnisschwierigkeiten zu kompensieren, die es durchaus gibt zwischen den unterschiedlichen Communities, auch in Europeana:

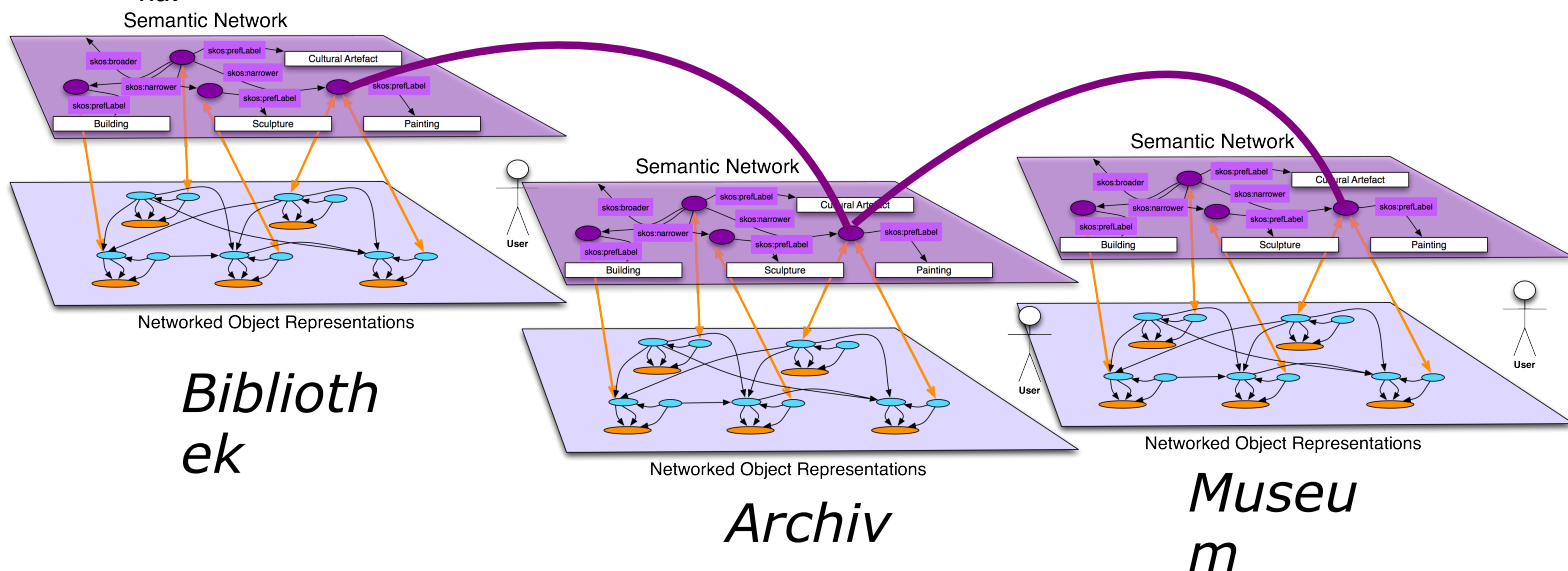


Abbildung 12: Verbindungen zwischen Kultur-Communities: Konzepte

Bibliotheken, Archive und Museen haben sehr unterschiedliche Auffassungen davon, was ein Objekt ist und was seine konstitutiven Eigenschaften eigentlich sind. Um diese Differenzen zu überbrücken, kann man die semantische Schicht verwenden, um eine Interoperabilität auf der Konzeptebene herzustellen. Und schließlich kann man dieses Modell einklinken in einen viel größeren Zusammenhang, der in der nachstehenden Abbildung sichtbar wird:

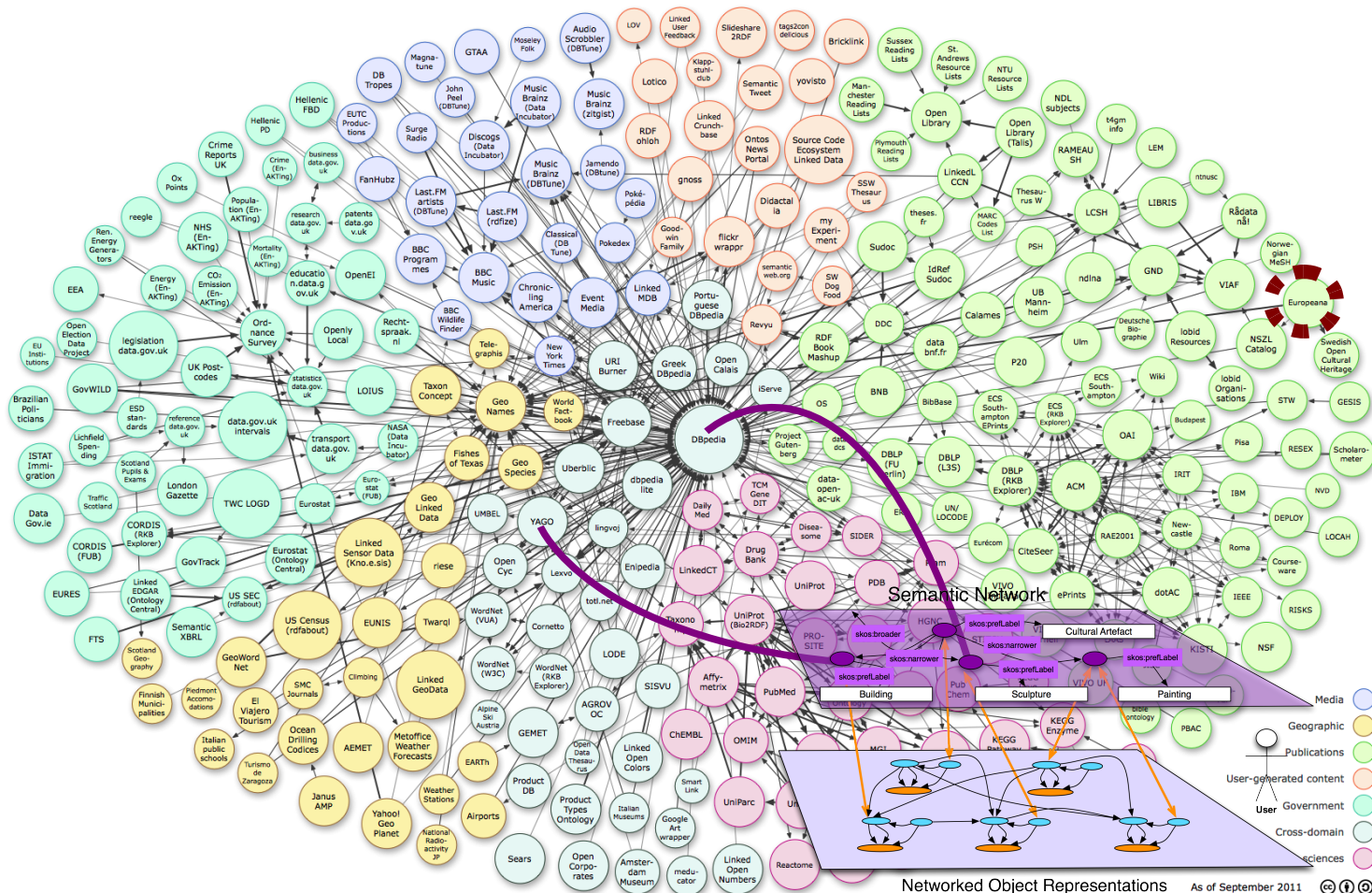


Abbildung 13: Europeana als Teil der LoD-Wolke

Es handelt sich um die Linked-Open-Data-Wolke mit der zentralen Ressource dbpedia, die ich vorhin schon erwähnt habe. Europeana ist seit September dieses Jahres Teil dieser Wolke, und technische Voraussetzung dafür war das EDM, denn alle Bestandteile dieser Wolke sind freie zugängliche RDF-Ressourcen und seitdem Europeana, zumindest in Gestalt des Linked-Data-Piloten, eine solche ist, werden wir Teil dieses Bildes.

Damit entstehen neue Fragen und neue Chancen. Einige dieser Fragen sind in der folgenden Abbildung angedeutet:

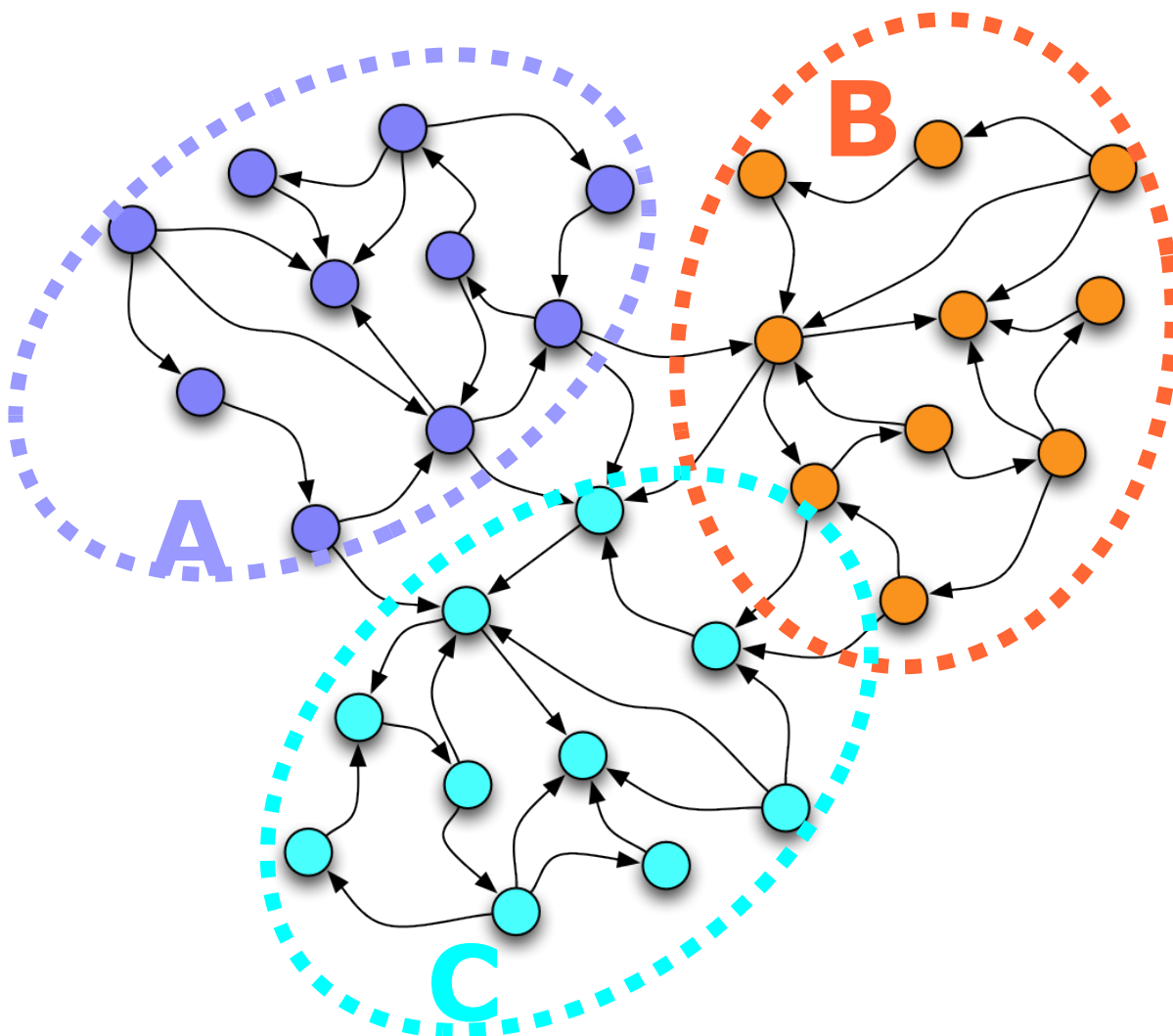


Abbildung 14: Neue Fragen

Jenseits der Katalogdatensätze im Kontext eines solchen graphenbasierten Aggregationsmodells kann man sich z.B. fragen, wo diese Aggregationen eigentlich beginnen, wo sie enden. Beziehungsweise – eine eng verwandte Frage – was konstituiert eigentlich die Grenzen der hier suggestiv farblich voneinander abgesetzten Aggregationen A, B und C? Wo hört das Objekt auf, wo beginnt sein Kontext – und wo endet dieser? Bis hin zu der folgenden, vielleicht anspruchsvollsten Frage: Welcher dieser Knoten war eigentlich zu einem gegebenen Zeitpunkt mit welchem anderen verbunden? Wie kann man das Modell also auch in der Zeit organisieren, wie kann man ein solches Konstrukt, solche Segmente des Linked Data Webs eigentlich versionieren? Wie kann man modellieren, welcher dieser Knoten wann Teil welcher Aggregation war? Es entstehen also damit eine ganze Menge neuer Herausforderungen, über die ich als Informationswissenschaftler natürlich glücklich bin, weil sie uns über Jahrzehnte in Lohn und Brot halten werden.

Aber wir erhalten auch neues Potenzial: Wir können Inferenzen produzieren, einfache, deterministische Schlußsysteme über Linked Data Ressourcen aufbauen, wie im folgenden Beispiel zu sehen:

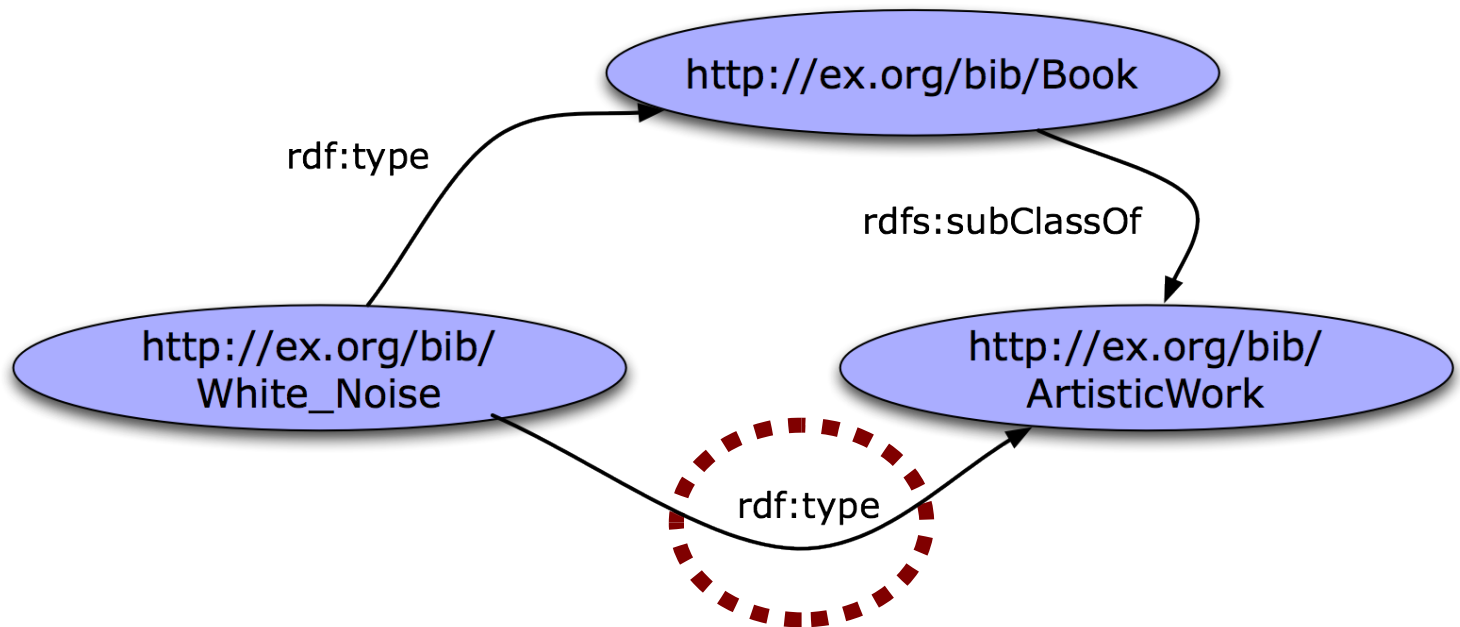


Abbildung 15: Neue Optionen: Inferenz

Gegeben eine Aussage, derzufolge eine Resource „White Noise“ vom `rdf:type` Buch ist. Ein zweites Tripel sagt aus, dass 'Buch' eine Unterklasse von 'Artistic Work' ist. Mithilfe dieser beiden Tripel kann eine Maschine - und dazu braucht es jetzt kein menschliches Zutun mehr - eine dritte Aussage inferieren, nämlich dass "White Noise" vom Typ Artistic Work sein muss.

Solche Verfahren haben bei entsprechender Skalierung ein unglaubliches Potenzial, gerade für Kultureinrichtungen. Das will ich Ihnen an einem leicht humoristischen Beispiel erläutern das ich von Jill Cousins, der Leiterin des Europeana Office übernommen habe.

Hier haben wir ein Aussagensystem über Wikinger:

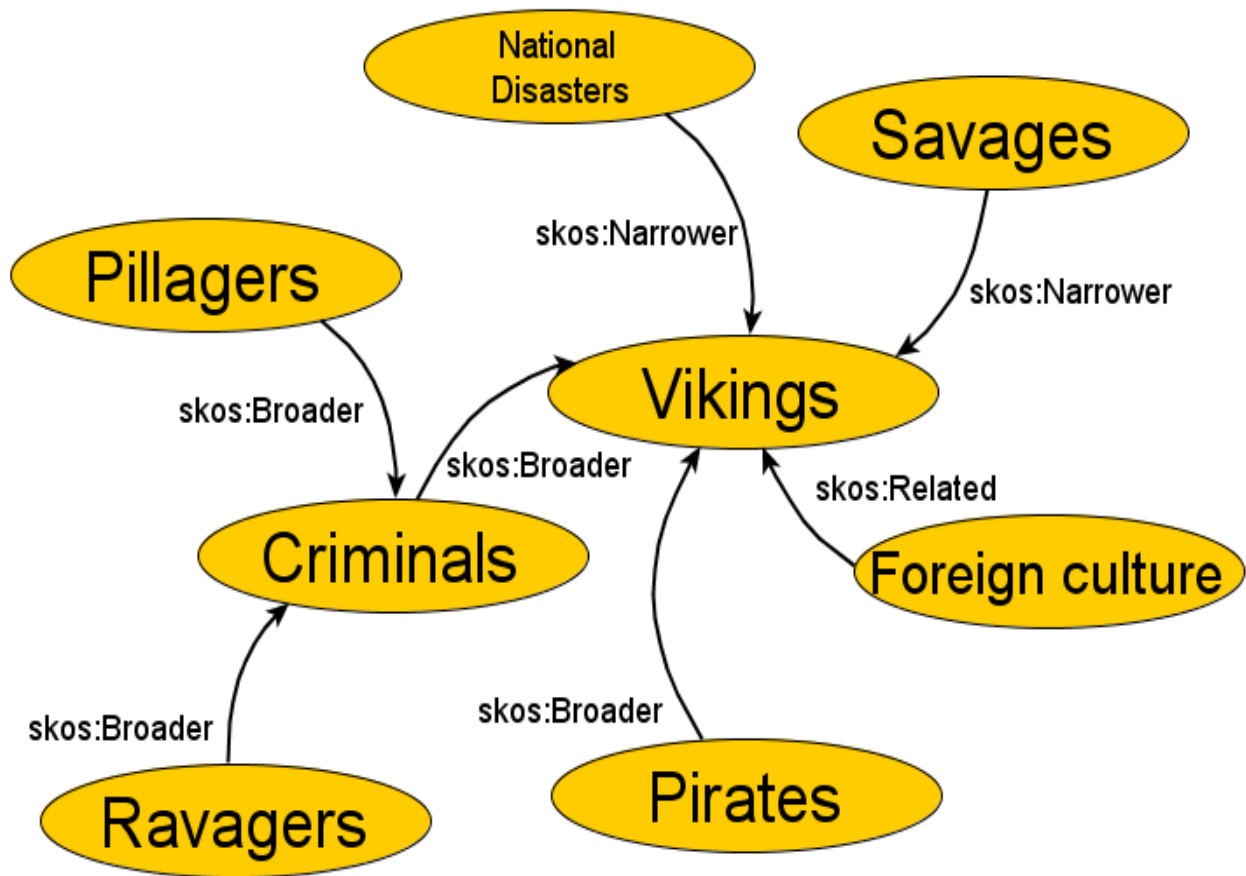


Abbildung 16: 'Böse' Wikinger

Es ist ganz offenkundig nicht aus der norwegischen Perspektive formuliert. Da ist die Rede von „Criminals“ und „Pillagers“, „National Disasters“, „Savages“, „Foreign Culture Pirates“ und „Ravagers“ - sämtlich Tripel, die hier in SKOS typisiert Aussagen über die Wikinger machen.

Dazu gibt es ein zweites, diesmal aus norwegischer Perspektive formuliertes Aussagensystem:

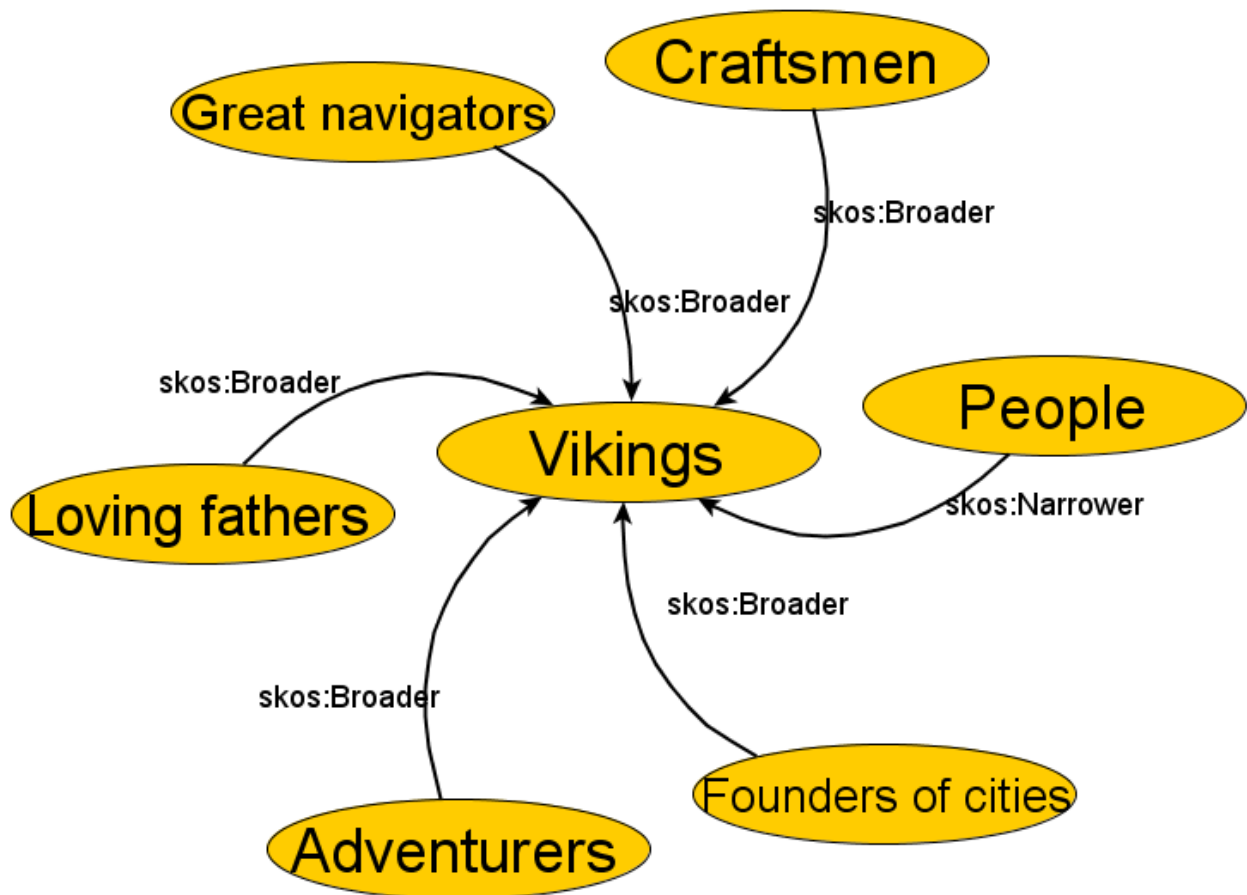


Abbildung 17: 'Gute' Wikinger

Hier haben wir es nun zu tun mit "Great Navigators", "Craftsmen", "People", "Founders of Cities", "Adventurers" und "Loving Fathers". Schalten Sie jetzt die beiden Aussagensysteme zu sammen, so erhalten Sie das folgende Bild:

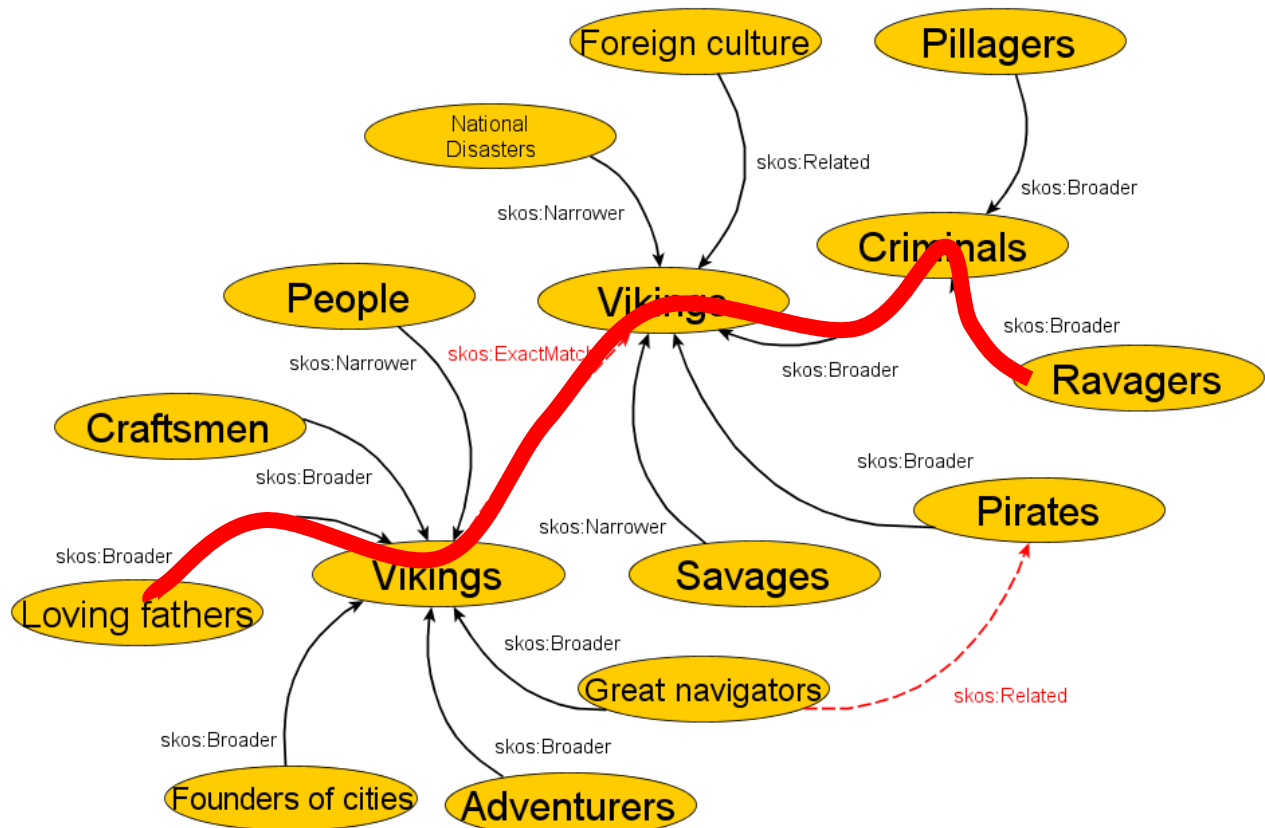


Abbildung 18: ... und die Kombination beider Aussagensysteme

Indem wir zwischen den beiden zentralen „Vikings“-Knoten die Beziehung *skos:Exact Match* herstelle ermögli­che ich den Aufbau eines komplexeren, durch die rote Linie angedeu­ten komplexeren Graphen, der von den „Loving Fathers“ bis zu den „Ravagers“ reicht.

Man kann auf diese Weise in solchen Aussagensystemen sehr komplexe Zusammenhänge herstellen, man kann letztlich Information über Kulturerbe, die so modelliert ist, nicht nur re­präsen­tieren, sondern auch prozessieren. Das kann mit maschineller Hilfe geschehen, und damit werden auch wissenschaftlich interessante Operationen möglich, die in der Vergangenheit - je denfalls in dieser Skalierung - nicht möglich waren.

Damit komme ich zum Schluss, zu den Potenzialen und Fragen für Kultureinrichtungen. Wenn Sie sich all das, was ich Ihnen jetzt vorzustellen versucht habe, durch den Kopf gehen las­sen, dann wird klar, dass wir einiges gewinnen. Wir können zeitlichen, räumlichen, begrifflichen Kontext von Artefakten in ganz anderer Weise modellieren, als es in der Vergangenheit der Fall gewesen ist, auch mit einer ganz anderen Dynamik als das bislang möglich war. Wir können konkurrierende Aussagensysteme zulassen. Wir erhalten ein ganz neues Interaktionspotenzial - auch mit unseren Nutzern, denn solche Mechanismen kann man um soziale Interaktionsmecha­nismen ergänzen, und dann können Nutzer von Kultureinrichtungen mit deren Aussagensystem interagieren, es erweitern. Damit werden aus den Objektsilos, die viele unserer Kultureinrich­tun­gen jedenfalls lange Jahre vornehmlich gewesen sind, Systeme für die Wissensgenerierung, die dann letztlich ein ganz anderes wissenschaftliches Potenzial haben. Gedächtniseinrichtungen werden damit Teil globaler Wissensarchitektur.

Das waren die attraktiven Aspekte dieser Entwicklung - hinter all dem steckt aber auch eine weniger angenehme Frage. Nämlich: wie viele partikuläre Institutionen brauchen wir unter diesen Bedingungen denn tatsächlich noch? Wie viel Institutionenbezug löst sich in den virtuel­len Link-Beziehungen im Web auf? Wie viel harten Institutionenbezug werden wir wirklich erhal

ten müssen? Das sind Fragen, über die wir in den nächsten Jahren intensiv nachdenken sollten, und zwar bevor irgendwann Politiker auf die Idee kommen, dass man diese ganzen Institutionen einfach nicht mehr braucht, weil sie ja alle im Linked Data Web gut aufgehoben sind.

In diesem Sinne danke ich Ihnen für Ihre hoffentlich entstandene Nachdenklichkeit und eventuelle Fragen.